AMENDMENT

In the Specification

Please replace the paragraph on page 6, lines 4-9 with the following:

This system and method employs the use of gene expression microarrays. For example, microarrays consisting of full length genes or gene fragments on a substrate may be formed. These arrays can then be tested with samples treated with a substances to elucidate the gene expression pattern associated with treatment with the substance. This gene pattern can be compared with gene expression patterns of compounds associated with known toxicological responses.

Please replace the paragraph on page 8, lines 10-15 with the following:

In a further preferred aspect of this embodiment, the summary scores are subjected to logistic regression analysis, resulting in a predictive model. In this aspect of the embodiment, the input data are the summary scores per sample, which is an indicator for each sample; the analysis is a logistic regression analysis mapping the summary scores to a 0 to 1 scale of toxicity; and the output data are one or more mathematical formulae that convert a column of average differences into a single 0 to 1 toxicological score for a sample.

Please replace the paragraph on page 9, lines 1-3 with the following:

In correlating these other studies, one preferably compare gene lists for patterns of interest between studies of related compounds to arrive at a consensus set of genes involved in a toxicological response.

Please replace the paragraph on page 9, lines 4-10 with the following:

In another preferred embodiment of the present invention, the goal of the method for assessing the toxicity and toxicology of a substance is to use gene expression to predict whether a compound has a high probability of being toxic at a given dose. In this preferred embodiment, patterns of gene expression can be compared against known "toxic" patterns and a similarity score calculated. Preferably, the methodology associated with this preferred embodiment includes identification of gene expression patterns associated with toxicity; quantification of this association; development of a statistical inference of similarity; and validation of results.

Please replace the paragraph on page 9, lines 11-13 with the following:

It will be appreciated that in such a modeling, there can be a number of different types of markers, including general markers, group markers (for example, cholestasis, necrosis, stenosis), and compound specific markers.

Please replace the paragraph on page 9, lines 20-24 with the following:

In another preferred embodiment of the present invention, there are various stages of model development. These preferably include: selection (determination of relevant expression patterns that are time stable and dose dependent); quantification (production of composite measures that define patterns); prediction (use of composite measures to assign probability of

Please replace the paragraph on page 22, lines 18-23 with the following:

The term "mismatch control" refers to a probe that has a sequence deliberately selected not to be perfectly complementary to a particular target sequence. The mismatch

control typically has a corresponding test probe that is perfectly complementary to the same particular target sequence. The mismatch may comprise one or more bases: While the mismatch(s) may be located anywhere in the mismatch probe, terminal mismatches are less desirable as a terminal mismatch is less likely to prevent hybridization of the target sequence. In a particularly

Please replace the paragraph on page 27, lines 4-9 with the following:

For example if a single doses of a drug and a vehicle is administered at three time points. Then, for each time point a gene would demonstrate a basic pattern of either upregulated, downregulated, or not significantly changing. The number of patterns produced would then be three for each time which would mean that 3 x 3 x 3-= 27 patterns can be produced. When one has multiple doses and a larger number of time points, the number of patterns can be extensive. But only a small number of these patterns are useful.

Please replace the paragraph on page 28, lines 3-6 with the following:

With regard to quantification of the toxicological response, principal component analysis (PCA) is employed. As an input, genes are selected for patterns that are biologically relevant to the toxicological process. Then, PCA analysis is performed on all samples. The resultant output is 1 to 8 summary scores for each sample.

Please replace the paragraph on page 29, lines 1-11 with the following:

multidimensional scaling, clustering, and neural networks. A general discussion of each technique can be found in "Multivariate Analysis, Prentice Hall ISBN 0-13-894858," which is incorporated herein by reference. All of these methods work by making composite measures from the many measurements taken from each object. With gene expression patterns there are several time and dose points which represent multiple objects that are

grouped together. None of these techniques are sufficient alone to represent this order of complexity. Contrast analysis allows one to identify measurements that are partial independent of time because they are time stable yet are affected by toxic doses more then non toxic doses. The PCA combines these many measurements into a series of orthogonal composite measures. Since these composite measures are non correlated by definition the problem of multicolinearity which can decrease the power of logistic regression is eliminated. By combining these techniques in the order described many of the limitations of each individual technique is reduced

Please replace the paragraph on page 31, lines 1-4 with the following:

This model is used to predict the probability of toxicity for each of the J samples. If the probability for the known toxins is consistently high and the probability for the known non-toxins is consistently low, then the model is accepted. Otherwise, the gene selection criteria is altered, and the multivariate statistical analysis is repeated.

Please replace the paragraph on page 31, lines 5-14 with the following:

The invention consists of three distinct stages. At each stage, small variations in technique can be used to accomplish the same task. The first stage, selection of time stable and dose dependent patterns by contrast analysis, can be altered by changing the method of measuring variation. The method used is based on analysis of variance, where the time component and dose component are assessed simultaneously. One could use a series of t test on individual parts of the pattern to get a collective set of p values that could approximate our method of measuring variation. One could also set an arbitrary fractional cutoff, mean or median of experimental group divided by control group, to approximate the measurement of variation for each part of the pattern that is then use in the next to stages of analysis. The novel feature is to find time, stable and dose dependent patterns with a predicted p value for that pattern.